

#2

OIPE

RAW SEQUENCE LISTING

DATE: 12/06/2001

PATENT APPLICATION: US/09/990,385

TIME: 17:15:54

Input Set : A:\PTO.DC.txt

Output Set: N:\CRF3\12062001\I990385.raw

SEQUENCE LISTING

4 (1) GENERAL INFORMATION:
5 (i) APPLICANT: Koji YANAI et al.
6 (ii) TITLE OF INVENTION: BETA-FRUCTOFURANOSIDASE AND ITS GENE, METHOD OF
7 ISOLATING BETA-FRUCTOFURANOSIDASE GENE, SYSTEM FOR
PRODUCING
8 BETA-FRUCTOFURANOSIDASE, AND BETA-FRUCTOFURANOSIDASE
VARIANT
9 (iii) NUMBER OF SEQUENCES: 35
10 (iv) CORRESPONDENCE ADDRESS:
11 (A) ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
12 (B) STREET: 2033 K Street, N.W., Suite 800
13 (C) CITY: Washington
14 (D) STATE: D.C.
15 (E) COUNTRY: U.S.A.
16 (F) ZIP: 20006
17 (v) COMPUTER READABLE FORM:
18 (A) MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
19 (B) COMPUTER: IBM Compatible
20 (C) OPERATING SYSTEM: MS-DOS
21 (D) SOFTWARE: Wordperfect 5.1
22 (vi) CURRENT APPLICATION DATA:
C--> 23 (A) APPLICATION NUMBER: US/09/990,385
C--> 24 (B) FILING DATE: 10-Sep-1998
25 (vii) PRIOR APPLICATION DATA:
26 (A) APPLICATION NUMBER: 09/142,623
27 (B) FILING DATE: September 10, 1998
28 (viii) ATTORNEY/AGENT INFORMATION:
29 (A) NAME: Lee Cheng
30 (B) REGISTRATION NUMBER: 40,949
31 (C) REFERENCE/DOCKET NUMBER: 2001-1611
32 (ix) TELECOMMUNICATION INFORMATION:
33 (A) TELEPHONE: 202-721-8200
34 (B) TELEFAX: 202-721-8250
35 (2) INFORMATION FOR SEQ ID NO: 1:
36 (i) SEQUENCE CHARACTERISTICS:
37 (A) LENGTH: 635 amino acid residues
38 (B) TYPE: amino acid
39 (C) STRANDEDNESS: Not relevant
W--> 40 (D) TOPOLOGY: Not relevant
41 (ii) MOLECULE TYPE: protein
42 (vi) ORIGINAL SOURCE:
43 (A) ORGANISM: Microorganism: Aspergillus niger ACE-2-1
44 (ATCC 20611)
45 (ix) FEATURE:
46 (A) NAME/KEY: mat peptide
47 (B) LOCATION: 1..635
48 (C) IDENTIFICATION METHOD: E

49 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/990,385

DATE: 12/06/2001

TIME: 17:15:54

Input Set : A:\PTO.DC.txt

Output Set: N:\CRF3\12062001\I990385.raw

```

50 Ser Tyr His Leu Asp Thr Thr Ala Pro Pro Pro Thr Asn Leu Ser Thr
51 1 5 10 15
52 Leu Pro Asn Asn Thr Leu Phe His Val Trp Arg Pro Arg Ala His Ile
53 20 25 30
54 Leu Pro Ala Glu Gly Gln Ile Gly Asp Pro Cys Ala His Tyr Thr Asp
55 35 40 45
56 Pro Ser Thr Gly Leu Phe His Val Gly Phe Leu His Asp Gly Asp Gly
57 50 55 60
58 Ile Ala Gly Ala Thr Thr Ala Asn Leu Ala Thr Tyr Thr Asp Thr Ser
59 65 70 75 80
60 Asp Asn Gly Ser Phe Leu Ile Gln Pro Gly Gly Lys Asn Asp Pro Val
61 85 90 95
62 Ala Val Phe Asp Gly Ala Val Ile Pro Val Gly Val Asn Asn Thr Pro
63 100 105 110
64 Thr Leu Leu Tyr Thr Ser Val Ser Phe Leu Pro Ile His Trp Ser Ile
65 115 120 125
66 Pro Tyr Thr Arg Gly Ser Glu Thr Gln Ser Leu Ala Val Ala Arg Asp
67 130 135 140
68 Gly Gly Arg Arg Phe Asp Lys Leu Asp Gln Gly Pro Val Ile Ala Asp
69 145 150 155 160
70 His Pro Phe Ala Val Asp Val Thr Ala Phe Arg Asp Pro Phe Val Phe
71 165 170 175
72 Arg Ser Ala Lys Leu Asp Val Leu Leu Ser Leu Asp Glu Glu Val Ala
73 180 185 190
74 Arg Asn Glu Thr Ala Val Gln Gln Ala Val Asp Gly Trp Thr Glu Lys
75 195 200 205
76 Asn Ala Pro Trp Tyr Val Ala Val Ser Gly Gly Val His Gly Val Gly
77 210 215 220
78 Pro Ala Gln Phe Leu Tyr Arg Gln Asn Gly Gly Asn Ala Ser Glu Phe
79 225 230 235 240
80 Gln Tyr Trp Glu Tyr Leu Gly Glu Trp Trp Gln Glu Ala Thr Asn Ser
81 245 250 255
82 Ser Trp Gly Asp Glu Gly Thr Trp Ala Gly Arg Trp Gly Phe Asn Phe
83 260 265 270
84 Glu Thr Gly Asn Val Leu Phe Leu Thr Glu Glu Gly His Asp Pro Gln
85 275 280 285
86 Thr Gly Glu Val Phe Val Thr Leu Gly Thr Glu Gly Ser Gly Leu Pro
87 290 295 300
88 Ile Val Pro Gln Val Ser Ser Ile His Asp Met Leu Trp Ala Ala Gly
89 305 310 315 320
90 Glu Val Gly Val Gly Ser Glu Gln Glu Gly Ala Lys Val Glu Phe Ser
91 325 330 335
92 Pro Ser Met Ala Gly Phe Leu Asp Trp Gly Phe Ser Ala Tyr Ala Ala
93 340 345 350
94 Ala Gly Lys Val Leu Pro Ala Ser Ser Ala Val Ser Lys Thr Ser Gly
95 355 360 365
96 Val Glu Val Asp Arg Tyr Val Ser Phe Val Trp Leu Thr Gly Asp Gln
97 370 375 380
98 Tyr Glu Gln Ala Asp Gly Phe Pro Thr Ala Gln Gln Gly Trp Thr Gly

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/990,385

DATE: 12/06/2001

TIME: 17:15:54

Input Set : A:\PTO.DC.txt

Output Set: N:\CRF3\12062001\I990385.raw

```

99 385          390          395          400
100 Ser Leu Leu Leu Pro Arg Glu Leu Lys Val Gln Thr Val Glu Asn Val
101          405          410          415
102 Val Asp Asn Glu Leu Val Arg Glu Glu Gly Val Ser Trp Val Val Gly
103          420          425          430
104 Glu Ser Asp Asn Gln Thr Ala Arg Leu Arg Thr Leu Gly Ile Thr Ile
105          435          440          445
106 Ala Arg Glu Thr Lys Ala Ala Leu Leu Ala Asn Gly Ser Val Thr Ala
107          450          455          460
108 Glu Glu Asp Arg Thr Leu Gln Thr Ala Ala Val Val Pro Phe Ala Gln
109 465          470          475          480
110 Ser Pro Ser Ser Lys Phe Phe Val Leu Thr Ala Gln Leu Glu Phe Pro
111          485          490          495
112 Ala Ser Ala Arg Ser Ser Pro Leu Gln Ser Gly Phe Glu Ile Leu Ala
113          500          505          510
114 Ser Glu Leu Glu Arg Thr Ala Ile Tyr Tyr Gln Phe Ser Asn Glu Ser
115          515          520          525
116 Leu Val Val Asp Arg Ser Gln Thr Ser Ala Ala Ala Pro Thr Asn Pro
117          530          535          540
118 Gly Leu Asp Ser Phe Thr Glu Ser Gly Lys Leu Arg Leu Phe Asp Val
119 545          550          555          560
120 Ile Glu Asn Gly Gln Glu Gln Val Glu Thr Leu Asp Leu Thr Val Val
121          565          570          575
122 Val Asp Asn Ala Val Val Glu Val Tyr Ala Asn Gly Arg Phe Ala Leu
123          580          585          590
124 Ser Thr Trp Ala Arg Ser Trp Tyr Asp Asn Ser Thr Gln Ile Arg Phe
125          595          600          605
126 Phe His Asn Gly Glu Gly Glu Val Gln Phe Arg Asn Val Ser Val Ser
127          610          615          620
128 Glu Gly Leu Tyr Asn Ala Trp Pro Glu Arg Asn
129 625          630          635
132 (2) INFORMATION FOR SEQ ID NO: 2:
133   (i) SEQUENCE CHARACTERISTICS:
134       (A) LENGTH: 1905 base pairs
135       (B) TYPE: Nucleic acid
136       (C) STRANDEDNESS: Double stranded
137       (D) TOPOLOGY: Linear
138   (ii) MOLECULE TYPE: Genomic DNA
139   (vi) ORIGINAL SOURCE:
140       (A) ORGANISM: Microorganism: Aspergillus niger ACE-2-1
141                   (ATCC 20611)
142   (ix) FEATURE:
143       (A) NAME/KEY: mat peptide
144       (B) LOCATION: 1 .. 1905
145       (C) IDENTIFICATION METHOD: E
146   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
147 TCATACCACC TGGACACCAC GGCCCCGCCG CCGACCAACC TCAGCACCCT CCCCAACAAC   60
148 ACCCTCTTCC ACGTGTGGCG GCCGCGCGCG CACATCCTGC CCGCCGAGGG CCAGATCGGC   120
149 GACCCCTGCG CGCACTACAC CGACCCATCC ACCGGCCTCT TCCACGTGGG GTTCCTGCAC   180

```

RAW SEQUENCE LISTING

DATE: 12/06/2001

PATENT APPLICATION: US/09/990,385

TIME: 17:15:54

Input Set : A:\PTO.DC.txt

Output Set: N:\CRF3\12062001\I990385.raw

```

150 GACGGGGACG GCATCGCGGG CGCCACCACG GCCAACCTGG CCACCTACAC CGATACCTCC 240
151 GATAACGGGA GCTTCCTGAT CCAGCCGGGC GGAAGAAGAC ACCCCGTCGC CGTGTTCGAC 300
152 GCGCCGTCGA TCCCCGTCGG CGTCAACAA ACCTCCACCT TACTCTACAC CTCCGTCTCC 360
153 TTCCTGCCCA TCCACTGGTC CATCCCTAC ACCCGCGGCA GCGAGACGCA GTCGTTGGCC 420
154 GTCGCGCGCG ACGGCGGGCC CCGCTTCGAC AAGCTCGACC AGGGCCCCGT CATCGCCGAC 480
155 CACCCCTTCG CCGTCGACGT CACCGCCTTC CGCGATCCGT TTGTCTTCCG CAGTGCCAAG 540
156 TTGGATGTGC TGCTGTCTGT GGATGAGGAG GTGGCGCGGA ATGAGACGGC CGTGCAGCAG 600
157 GCCGTCGATG GCTGGACCGA GAAGAAGGCC CCCTGGTATG TCGCGGTCTC TGGCGGGGTG 660
158 CACGGCGTCG GGCCCGCGCA GTTCCTCTAC CGCCAGAACG GCGGGAACGC TTCCGAGTTC 720
159 CAGTACTGGG AGTACCTCGG GGAGTGGTGG CAGGAGGCGA CCAACTCCAG CTGGGGCGAC 780
160 GAGGGCACCT GGGCCGGGCG CTGGGGGTTC AACTTCGAGA CGGGGAATGT GCTCTTCCTC 840
161 ACCGAGGAGG GCCATGACCC CCAGACGGGC GAGGTGTTTCG TCACCCTCGG CACGGAGGGG 900
162 TCTGGCCTGC CAATCGTGCC GCAGGTCTCC AGTATCCACG ATATGCTGTG GGCGGCGGGT 960
163 GAGGTCGGGG TGGGCAGTGA GCAGGAGGGT GCCAAGGTCG AGTTCTCCCC CTCCATGGCC 1020
164 GGGTTTCTGG ACTGGGGGTT CAGCGCCTAC GCTGCGGCGG GCAAGGTGCT GCCGGCCAGC 1080
165 TCGGCGGTGT CGAAGACCAG CGGCGTGAGG GTGGATCGGT ATGTCTCGTT CGTCTGGTTG 1140
166 ACGGGCGACC AGTACGAGCA GGCGGACGGG TTCCCCACGG CCCAGCAGGG GTGGACGGGG 1200
167 TCGTGCTGTC TGCCGCGCGA GCTGAAGGTG CAGACGGTGG AGAACGTCGT CGACAACGAG 1260
168 CTGGTGCGCG AGGAGGGCGT GTCGTGGGTG GTGGGGGAGT CGGACAACCA GACGGCCAGG 1320
169 CTGCGCACGC TGGGGATCAC GATCGCCCGG GAGCCAAGG CGGCCCTGCT GGCCAACGGC 1380
170 TCGGTGACCG CGGAGGAGGA CCGCACGCTG CAGACGGCGG CCGTCGTGCC GTTCGCGCAA 1440
171 TCGCCGAGCT CCAAGTTCTT CGTGCTGACG GCCCAGCTGG AGTTCCCCGC GAGCGCGCGC 1500
172 TCGTCCCCGC TCCAGTCCGG GTTCGAAATC CTGGCGTCGG AGCTGGAGCG CACGGCCATC 1560
173 TACTACCACT TCAGCAACGA GTCGCTGGTC GTCGACCGCA GCCAGACTAG TCGGGCGGCG 1620
174 CCCACGAACC CCGGGCTGGA TAGCTTTACT GAGTCCGGCA AGTTGCGGTT GTTCGACGTG 1680
175 ATCGAGAACG GCCAGGAGCA GGTGAGACG TTGGATCTCA CTGTCGTCGT GGATAACCGG 1740
176 GTTGTCGAGG TGTATGCCAA CGGGCGCTTT GCGTTGAGCA CCTGGGCGAG ATCGTGGTAC 1800
177 GACAACCTCA CCCAGATCCG CTTCTTCCAC AACGGCGAGG GCGAGGTGCA GTTCAGGAAT 1860
178 GTCTCCGTGT CGGAGGGGCT CTATAACGCC TGGCCGGAGA GAAAT 1905

```

180 (2) INFORMATION FOR SEQ ID NO: 3:

181 (i) SEQUENCE CHARACTERISTICS:

182 (A) LENGTH: 20 amino acid residues

183 (B) TYPE: Amino acid

184 (C) STRANDEDNESS: Not relevant

185 (D) TOPOLOGY: Linear

186 (ii) MOLECULE TYPE: Peptide

187 (v) FRAGMENT TYPE: internal fragment

188 (vi) ORIGINAL SOURCE:

189 (A) ORGANISM: Microorganism: *Aspergillus niger* ACE-2-1

190 (ATCC 20611)

191 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

192 Leu Asp Gln Gly Pro Val Ile Ala Asp His Pro Phe Ala Val Asp Val

193 1 5 10 15

194 Thr Ala Phe Arg

195 20

197 (2) INFORMATION FOR SEQ ID NO: 4:

198 (i) SEQUENCE CHARACTERISTICS:

199 (A) LENGTH: 20 amino acid residues

200 (B) TYPE: Amino acid

RAW SEQUENCE LISTING

DATE: 12/06/2001

PATENT APPLICATION: US/09/990,385

TIME: 17:15:54

Input Set : A:\PTO.DC.txt

Output Set: N:\CRF3\12062001\I990385.raw

```

201          (C) STRANDEDNESS: Not relevant
202          (D) TOPOLOGY: Linear
203      (ii) MOLECULE TYPE: Peptide
204      (v) FRAGMENT TYPE: internal fragment
205      (vi) ORIGINAL SOURCE:
206          (A) ORGANISM: Microorganism: Aspergillus niger ACE-2-1
207                      (ATCC 20611)
208      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
209 Val Glu Phe Ser Pro Ser Met Ala Gly Phe Leu Asp Trp Gly Phe Ser
210  1          5          10          15
211 Ala Tyr Ala Ala
212          20
214 (2) INFORMATION FOR SEQ ID NO: 5:
215      (i) SEQUENCE CHARACTERISTICS:
216          (A) LENGTH: 20 amino acid residues
217          (B) TYPE: Amino acid
218          (C) STRANDEDNESS: Not relevant
219          (D) TOPOLOGY: Linear
220      (ii) MOLECULE TYPE: Peptide
221      (v) FRAGMENT TYPE: internal fragment
222      (vi) ORIGINAL SOURCE:
223          (A) ORGANISM: Microorganism: Aspergillus niger ACE-2-1
224                      (ATCC 20611)
225      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
226 Val Gln Thr Val Glu Asn Val Val Asp Asn Glu Leu Val Arg Glu Glu
227  1          5          10          15
228 Gly Val Ser Trp
229          20
231 (2) INFORMATION FOR SEQ ID NO: 6:
232      (i) SEQUENCE CHARACTERISTICS:
233          (A) LENGTH: 20 amino acid residues
234          (B) TYPE: Amino acid
235          (C) STRANDEDNESS: Not relevant
236          (D) TOPOLOGY: Linear
237      (ii) MOLECULE TYPE: Peptide
238      (v) FRAGMENT TYPE: internal fragment
239      (vi) ORIGINAL SOURCE:
240          (A) ORGANISM: Microorganism: Aspergillus niger ACE-2-1
241                      (ATCC 20611)
242      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
W--> 243 Ala Ala Leu Leu Ala Xaa Gly Ser Val Thr Ala Glu Glu Asp Arg Thr
244  1          5          10          15
245 Leu Gln Thr Ala
246          20
248 (2) INFORMATION FOR SEQ ID NO: 7:
249      (i) SEQUENCE CHARACTERISTICS:
250          (A) LENGTH: 6 amino acid residues
251          (B) TYPE: Amino acid
252          (C) STRANDEDNESS: Not relevant

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/990,385

DATE: 12/06/2001

TIME: 17:15:55

Input Set : A:\PTO.DC.txt

Output Set: N:\CRF3\12062001\I990385.raw

L:23 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:24 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:40 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=1
L:243 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6
L:269 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=8
L:279 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=9
L:356 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=11
L:487 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=13
L:621 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=15
L:631 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=16
L:641 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=17
L:651 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=18
L:661 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=19
L:671 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=20
L:681 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=21
L:691 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=22
L:701 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=23
L:711 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=24
L:721 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=25
L:731 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=26
L:741 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=27
L:751 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=28
L:761 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=29
L:771 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=30
L:781 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=31
L:791 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=32
L:801 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=33
L:811 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=34
L:821 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=35